

An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing

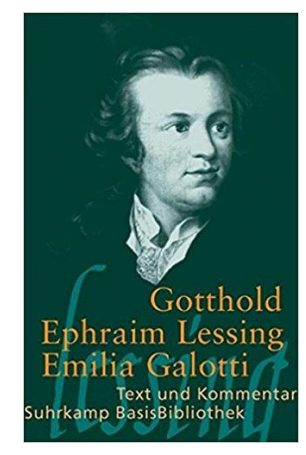
Thomas Schmidt, Media Informatics Group, Regensburg University
Manuel Burghardt, Computational Humanities Group, Leipzig University

Background

Context

Goal:
Performing Sentiment Analysis on historic German plays of G. E. Lessing

Data basis



- 12 plays by G.E. Lessing
- Written in the 18th century
- XML-annotated Files
- Structural information e.g. acts, scenes, speakers

Challenges

- No sentiment-annotated large scale corpora → limited machine learning possibilities
- Historical and poetic language
- Irony, sarcasm and lies

Lexicon-based methods

Sentiment lexicon

Wort	Polarität
1. Beschaeft	NEG-0.7
2. beherren	NEG-1
3. abkuehnen	NEG-1
4. Linnel	NEG-1
5. Erklaere	NEG-0.7
6. verachten	NEG-0.7
7. Schmeicheln	POS-0.1
8. abweisen	POS-1
9. gehoren	POS-1
10. Erloeb	POS-1
11. Proben	NEG-1
12. dratsch	NEG-1
13. Zankgerang	NEG-0.7
14. Stauernahme	POS-0.7
15. Hosenang	NEG-1
16. goll	POS-0.7
17. bedackeln	NEG-1

Sentiment bearing word
Polarity
Polarity strengths

Calculation

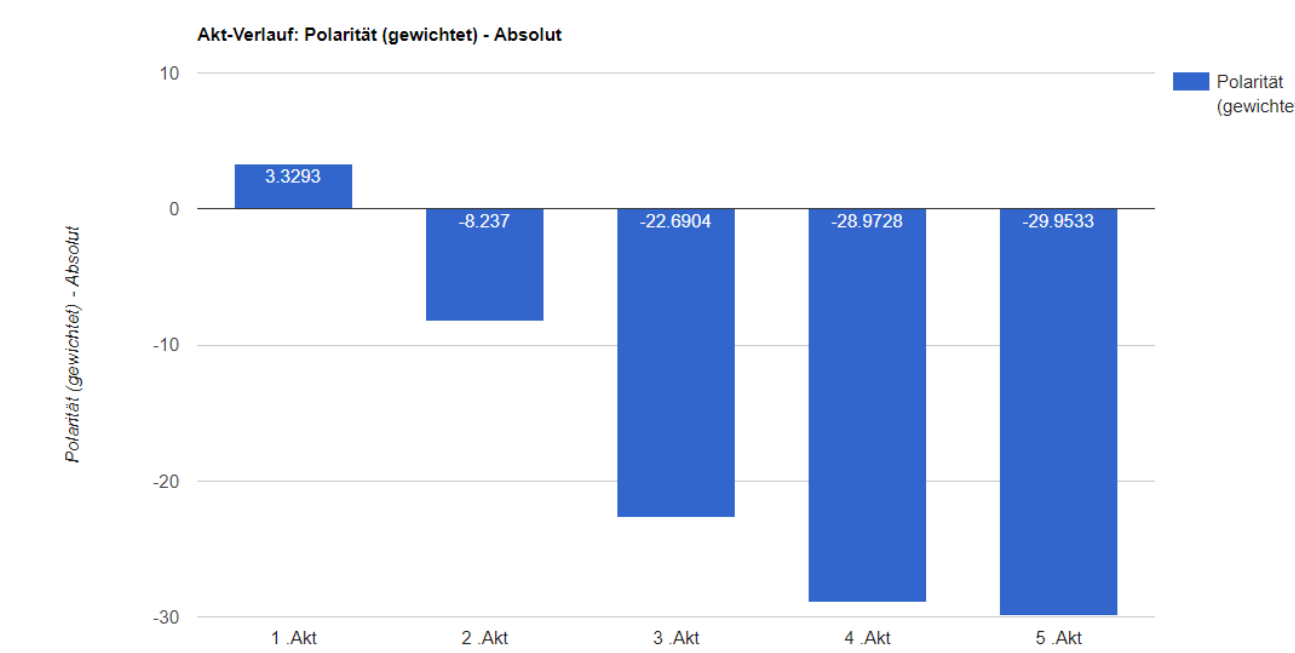
Ihre Zunge **verschont** nichts, auch dasjenige nicht, was ihr das **Heiligste** von der Welt sein sollte. **Pflicht**, **Tugend**, **Anständigkeit**, Religion: alles ist ihrem **Spotte** ausgesetzt. --

Final polarity

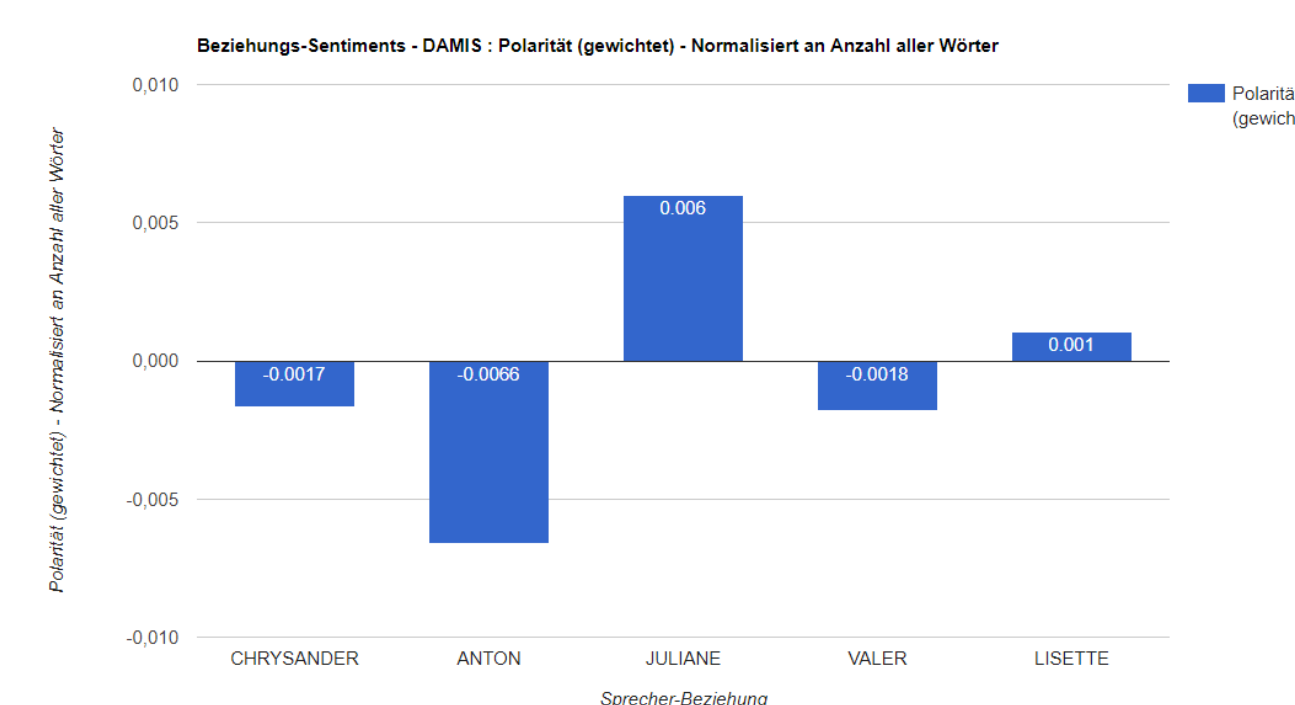
Positive: +4
Negative: -2
Overall-Polarity: +4 - 2 = +2 → Positive

Tool

Web-based tool to explore sentiment/emotion distributions and progressions
Online available: <https://tinyurl.com/y87ynedl>



Polarity Progression throughout all acts of the play *Emilia Galotti*



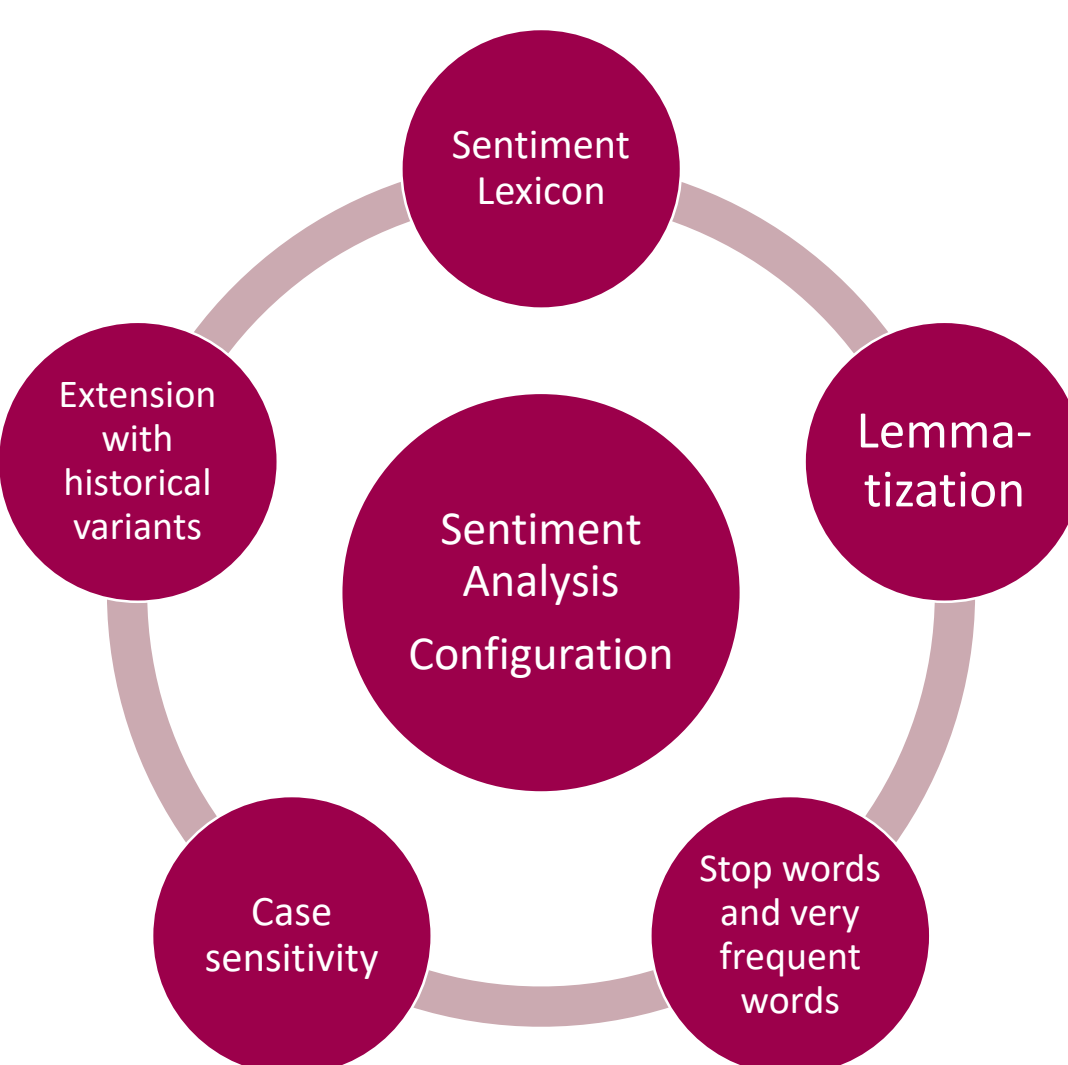
Character-to-Character polarity relations between DAMIS and the other characters of the play *Der junge Gelehrte*

Evaluation & Results

Methods

Research question:

- How well do German lexicon-based approaches and standard NLP-techniques work for Sentiment Analysis in this area?
- What is the best configuration of lexicon and NLP-techniques?



Approaches:

- 5 general purpose sentiment lexicons for German and a combined version of all lexicons
- 2 standard German lemmatizers
- Extension of the lexicons with historical linguistic variants (by a tool of the *Deutsches Textarchiv*)
- 4 different stop words lists
- Case sensitivity

Goal:

Evaluation and in-depth analysis of all combinations of approaches to identify best practices

Gold Standard creation:

- 200 representative speeches of the Lessing Corpus
- 5 annotators

Annotation scheme:

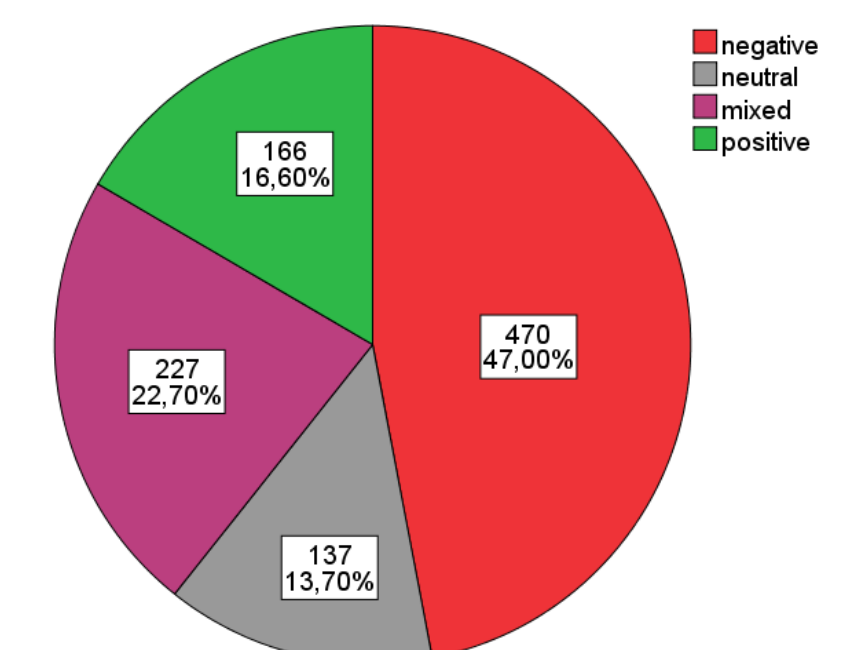
ADRAST:
Die fordern Sie nur von sich selbst.

JOHANN:
Hier muß ich hetzen. -- Ja, ja, Herr Theophan! es ist schon bekannt, daß Ihnen mein Herr ein Dorn in den Augen ist.

THEOPHAN:
Adrast, haben Sie es ihm befohlen, an Ihrer Stelle zu antworten?

Very Negative	Negative	Neutral	Mixed	Positive	Very Positive
	Negative			Positive	

Annotation distribution:



Final annotation:
Based on majority decision of binary polarity → 139 negative and 61 positive speeches

Levels of agreement:

Annotation	Krippendorff's α	Percentage of agreement
Polarity differentiated	0.22	40%
Binary polarity	0.47	77%

For more information about the annotation study see:
Schmidt, T., Burghardt, M. & Dennerlein, K. (2018). Sentiment Annotation of Historic German Plays: An Empirical Study on Annotation Behavior. In: Sandra Kübler, Heike Zinsmeister (eds.), *Proceedings of the Workshop on Annotation in Digital Humanities (annDH 2018)* (pp. 47-52). Sofia, Bulgaria. Retrieved from <http://ceur-ws.org/Vol-2155/schmidt.pdf>

Results

Evaluation of over 400 different configurations of sentiment lexicons and NLP-techniques

Metrics:

accuracy, precision, recall and F-measure

Wort	CTF-Messung	Lexikon	Stopwörter	CaseSensitivity	Ergebnis	F-Messung
1. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
4. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
6. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
7. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
10. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
11. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
12. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
13. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
14. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
15. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
16. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
17. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
18. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
19. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
20. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
21. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
22. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
23. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
24. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
26. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
27. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
28. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
29. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
30. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
31. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
32. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
33. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
34. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
35. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
36. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
37. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
38. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
39. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
40. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
41. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
42. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
43. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
44. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
45. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
46. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
47. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
48. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
49. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50. jenseits	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

Table snippet of the results of the evaluation of all configurations

Best configuration:

- Sentiment lexicon *SentiWortschatz*
 - Extension with historical lexical variants
 - Lemmatization via *treetagger*
 - No stop words lists
 - Ignoring case sensitivity
- Accuracy: 0.705
Average F-measure: 0.63
141 speeches are correctly predicted
- Above random baseline and slightly above majority baseline

Main Findings:

- Best lexicons/dictionaries: *SentiWortschatz* and the combined lexicon
- Lexicons with polarity strengths (e.g. polarity on a continuous scale from -1 to +1) perform better than lexicons with dichotomous polarity annotations
- Extension with historical linguistic variants consistently yields the strongest performance boost
- Stop words lists have differentiated influences based on the used lexicon. Some lexicons consist of stop words or very frequent words. Therefore stop words lists are recommended to pretend the sentiment calculation to be skewed
- Lemmatizers have problems with the historical language and vocabulary but perform equally good

Future plans:

- Exploring ML, hybrid sentiment analysis approaches and domain specific sentiment analysis
- More large scale annotation studies
- Broaden the scope, e.g. different authors, genres and eras
- Gathering more insights about the requirements of literary scholars